What can x-ray scattering tell us about the radial distribution functions of water?

Jon M. Sorenson
Department of Chemistry, University of California, Berkeley, California 94720

Greg Hura
Graduate Group in Biophysics, University of California, Berkeley and Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Robert M. Glaeser
Department of Molecular and Cell Biology, University of California, Berkeley and Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Teresa Head-Gordon[a)]
Physical Biosciences & Life Sciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, California 94720

(Received 26 May 2000; accepted 30 August 2000)

We present an analysis of the Advanced Light Source (ALS) x-ray scattering experiment on pure liquid water at ambient temperature and pressure described in the preceding article. The present study discusses the extraction of radial distribution functions from the x-ray scattering of molecular fluids. It is proposed that the atomic scattering factors used to model water be modified to include the changes in the intramolecular electron distribution caused by chemical bonding effects. Based on this analysis we present a $g_{OO}(r)$ for water consistent with our recent experimental data gathered at the ALS, which differs in some aspects from the $g_{OO}(r)$ reported by other x-ray and neutron scattering experiments. Our $g_{OO}(r)$ exhibits a taller and sharper first peak, and systematic shifts in all peak positions to smaller r. Based on experimental uncertainties, we discuss what features of $g_{OO}(r)$ should be reproduced by classical simulations of nonpolarizable and polarizable water models, as well as *ab initio* simulations of water, at ambient conditions. We directly compare many water models and simulations to the present data, and discuss possible improvements in both classical and *ab initio* simulation approaches in the future. © *2000 American Institute of Physics.* [S0021-9606(00)50844-7]

INTRODUCTION

The structural investigation of water has strong historical precedence, tracing roots at least as far back as Roentgen's early work on the structure of water and the explanation of its density maximum,[1] and Bernal and Fowler's model of hydrogen-bonding structure in the liquid.[2] In principle, an accurate characterization of the molecular structure of liquid water can be found from solution scattering experiments. Narten and Levy's x-ray scattering studies[3] and neutron scattering experiments by Soper and Phillips[4] and Soper, Bruni, and Ricci[5] are commonly cited as the definitive sources for the radial distribution functions of the fluid at ambient conditions.

However, these experiments differ with respect to the oxygen–oxygen radial distribution function, $g_{OO}(r)$. Furthermore, the original neutron scattering experiments[4] in recent years have been revised from the earlier analysis first presented with the data.[5] This has led to uncertainty on the part of those developing or using water force fields as to the overall robustness of a given water model when comparing the simulated results to those of experiment.[6–9] In addition, the oxygen–oxygen radial distribution function for water is an important ingredient of many theories concerning processes in water including the treatment of orientational effects,[7,10] phase changes,[11] and hydrophobic hydration.[12–17] The question apparently remains, what can and cannot be said about the structure of water from scattering experiments? What similarities in the published radial distribution functions are to be taken rigorously, and what differences are due to experimental uncertainties? Better confidence in the experimental facts will hopefully also lead to a consolidation of the very large number of water models and simulation methods presently in use.

We have recently performed a new x-ray diffraction study of liquid water under ambient conditions at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory that takes advantage of various state-of-the-art features of a modern day experiment,[18] including quantitative characterization of the x-ray source together with the use of a more sophisticated CCD area detector. As shown in our preceding companion article, we believe that the differences seen between our x-ray intensity profile and previously reported scattering curves are significant.

The present article aims to make clearer the current experimental picture concerning the structure of liquid water as

a)Author to whom correspondence should be addressed.

probed by x-ray and neutron scattering experiments, addressing what features of $g_{OO}(r)$ should be reproduced by a water model, and which water models favorably reproduce those features. The theory section discusses issues surrounding the extraction of radial distribution functions from the x-ray scattering of molecular fluids. We evaluate current models of the molecular from factor for water and propose a modification of the standard approach used for analyzing x-ray scattering, motivated by the desire to include the changes in the intramolecular electron distribution caused by chemical bonding effects. In the results, we present a $g_{OO}(r)$ for water consistent with our recent experimental data gathered at the ALS, which is different than the $g_{OO}(r)$ reported by other x-ray and neutron scattering experiments. Most importantly, we discuss what features of $g_{OO}(r)$ should be reproduced by a classical or *ab initio* simulation of water at ambient conditions. In the discussion, we compare $g_{OO}(r)$ generated from simulations using many nonpolarizable and polarizable empirical force fields for water, as well as recently reported *ab initio* simulations, to our experimentally determined $g_{OO}(r)$. We conclude with a summary of our results and analysis, and examine how these new results impact current understanding of the structure of water.

THEORETICAL CONSIDERATIONS

X-ray scattering experiments probe the distribution of electron density in a sample through the well-established relationship in the first Born approximation,[19]

$$\frac{d\sigma}{d\Omega}=\left\langle\left|\sum_{k=1}^{N_e} b_e \exp(i\mathbf{Q}\cdot\mathbf{r}_k)\right|^2\right\rangle, \tag{1}$$

where $d\sigma/d\Omega$ is the differential scattering cross-section, the sum is over the N_e electrons in the sample, b_e is the scattering length for a single electron, the $\{\mathbf{r}_k\}$ are the positions of the electrons, and $\mathbf{Q}$ is the momentum transfer for the scattering process. In applying this formula it is required that various corrections to the experimental data have been made, accounting for effects such as incoherent scattering, beam polarization, multiple scattering, and container absorption.

It is convenient for analyzing the scattering from molecular liquids to further separate this formula into contributions from individual molecules (self-scattering) and intermolecular correlations. Furthermore, the assumption is commonly made that the scattering can be represented as arising from independent neutral atoms, each with a spherical electron density distribution. Within these approximations, the observed scattering from a molecular liquid can be written as

$$I(Q)=\sum_{ij} x_i x_j f_i(Q) f_j(Q)\frac{\sin Qr_{ij}}{Qr_{ij}} + \sum_{i\leqslant j} x_i x_j f_i(Q) f_j(Q) h_{ij}(Q), \tag{2}$$

where

$$h_{ij}(Q)=4\pi\rho\int_0^\infty r^2\,dr(g_{ij}(r)-1)\frac{\sin Qr}{Qr}. \tag{3}$$

The sums are over the atom types present in the sample, x_i is the atomic fraction of atom type i, $f_i(Q)$ is the atomic scattering factor[20] for atom type i, ρ is the atomic density, r_{ij} are the intramolecular distances between atom centers, and $g_{ij}(r)$ is the radial distribution function describing intermolecular correlations between atom types i and j.

The molecular form factor

The first term in Eq. (2) corresponds to intramolecular scattering arising from the electrons centered on the individual atoms comprising the water monomer. This term is referred to as the molecular form factor or $\langle F(Q)^2\rangle$. The approximation written above, where each atom is considered to have a spherical electron density distribution, is known as the Debye approximation.[21] This expression is valid for a rigid molecule framework; we can include vibrational effects on the bond lengths by the simple modification[3,22]

$$\langle F(Q)^2\rangle=\sum_{ij} f_i(Q) f_j(Q)\exp(-b_{ij}Q^2/2)\frac{\sin Qr_{ij}}{Qr_{ij}}, \tag{4}$$

where b_{ij} is the variance of the length of the bond between atoms i and j. When using this form it is important that the geometry and vibrational corrections for the monomer correspond to those appropriate for the liquid phase. However, even after correcting for the proper geometry of the molecule in the condensed phase, this approximation is flawed. Approximating the electron density distribution of water as a superposition of electron densities centered on the individual atoms is not faithful to the true charge distribution. The individual atomic scattering factors[18] are calculated for the isolated atoms; as the atoms come together and covalently bond to form water, the charge density is changed significantly, with a depletion of electron density from the hydrogens and more electron density centered on the oxygen.[23,8]

In principle we can calculate the deviation from Eq. (4) using better approximations to the electron density of the monomer derived from (gas phase) *ab initio* calculations. The most commonly used molecular form factor for water derives from calculations by Blum[24] from self-consistent field-molecular orbital (SCF-MO) wave functions calculated by Moccia in 1964.[25] His curve for $\langle F(Q)^2\rangle$ as well as the Debye approximation result are shown in Fig. 1. This curve has been recently recalculated using quantum chemical calculations that include electron correlation and a significantly larger basis set.[26] The resulting curve, also shown in Fig. 1, does not differ significantly from the SCF calculation. Also shown in Fig. 1 is the experimentally derived $\langle F(Q)^2\rangle$ from recent x-ray scattering studies of gas-phase water after subtraction of incoherent scattering;[27] the experimental result agrees excellently with the calculated curves. We therefore can conclude that the spherically averaged Fourier transform of the electron density of the water monomer is relatively insensitive to the specific quality of the underlying wave functions. It is important to note that all three of these curves differ significantly from the curve resulting from the Debye approximation in the region of 1.0–4.0 Å^{-1}, which corresponds to the region of most interest for extracting the intermolecular correlations of water.

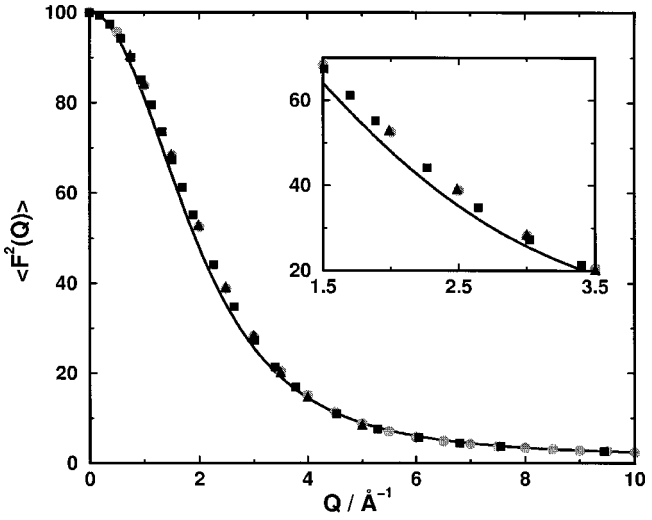

FIG. 1. $\langle F(Q)^2 \rangle$ for water in the gas-phase. Legend: Debye approximation, Eq. (4) (solid line); SCF calculation on water monomer (Ref. 24) (circles); CI calculation on water monomer (Ref. 26) (squares); experiment (Ref. 27) (triangles).

While we can state with certainty which is the proper $\langle F(Q)^2 \rangle$ for the water monomer in the gas phase, the intramolecular term in Eq. (2) should correspond to the electron distribution around the water monomer in the condensed phase. From the simple observation that the dipole moment of water in the gas phase[28] is 1.86 D while it is estimated to increase to 2.6–3.0 D in the liquid phase,[8,29] it is apparent that the monomer charge distribution changes significantly upon solvation. Again we can quantify this change with theoretical calculations. The most recent *ab initio* simulation studies of water[8] report that the electron distribution around a single water molecule is much changed from the gas phase, with more charge residing on the oxygen and a more spherical distribution of charge. An attempt has been made to include these types of observations into an analysis of the x-ray scattering from water.[23] While the change in the predicted intramolecular scattering was found to be significant, it had little effect on the observed intermolecular correlations because of statistical cancellations.[23] In our present work we have used the most accurately known gas-phase data on the molecular form factor[26] for normalizing and modeling the observed scattering. We will discuss below ways this approach can be augmented and their consequent effects.

The radial distribution functions

After the molecular form factor has been accounted for in Eq. (2), we can, in principle, extract the intermolecular correlations from the second term in this equation. It has often been stated that the x-ray scattering of water probes almost entirely the oxygen–oxygen correlations only. While this statement seems self-evident given the scattering power of oxygen relative to hydrogen, the reason for this has not been clearly enunciated. Egelstaff and Root have suggested[30] that the weight of $x_O^2 f_O^2$ vs. the weight of $x_O x_H f_O f_H$ for the intermolecular scattering terms is only 2:1 because of the increased number of O–H pairs in the liquid vs. O–O pairs. So it would appear that O–H correlations contribute as much as 33% of the observed scattering, and the molecular centers radial distribution functions reported by many experimental groups[3,31] should not be interpreted as $g_{OO}(r)$, as is commonly seen in the computational literature. In fact, this argument rests on the flawed assumption that charge distributions around each atom in the molecule correspond to the spherical electron density of the isolated atom. As emphasized above, the results of various investigations of the dipole moment of bulk water[6,8,29,9,32] indicate that the water molecule is much better represented as having far more charge around the oxygen atom than is the case for the isolated atom.

To illustrate the effect of this, consider the following simple approximation. If we redistribute charge on the water monomer by placing an extra 4/3 units of charge on the oxygen atom and remove 2/3 units of charge from each of the hydrogen atoms, we can augment the dipole moment to be 2.8 D, more typical of the condensed phase. When these considerations are placed into Egelstaff's argument, we find that the weight of O–O terms vs O–H terms in the observed scattering is actually 7:1, giving support to the picture that the centers radial distribution is mostly comprised of oxygen–oxygen correlations.

These considerations highlight the difficulty of calculating the predicted x-ray scattering for water. All applications of Eq. (2) in the water scattering literature have always assumed spherical atom electronic distributions using the atomic scattering factors derived from isolated atoms. This has the effect of overly weighting the O–H correlations in the predictions and further obscuring whether or not a simulated or extracted $g(r)$ corresponds to reality.

Modified atomic scattering factors

Deviations from the independent atom model (IAM) due to chemical bonding effects have been known and studied for many years.[33] Coppens and others have developed schemes for modifying scattering factors to take account of chemical bonding, including the contraction or dilation of valence electron density,[34] and allowing for aspherical atomic scattering factors.[35] However, to our knowledge these ideas have remained in the inorganic small molecule diffraction literature and have never found expression in the analysis of liquid-state x-ray scattering.

As discussed above, the Debye approximation performs inadequately for $\langle F(Q)^2 \rangle$ for gas phase water using the standard atomic scattering factors.[20] A simple modification is to scale the atomic scattering factors by the proper factor which gives a value of 1.86 D for the dipole moment of gas-phase water, i.e., multiply $f_O(Q)$ by 1.11 and $f_H(Q)$ by 0.56, and recompute $\langle F(Q)^2 \rangle$ using Eq. (4). The result of this calculation is compared to *ab initio* calculations in Fig. 2. We can see that this simple adjustment has greatly improved agreement at small Q, but at the sacrifice of agreement at large Q. The reason for this is apparent; the large Q tails of the atomic scattering factors probe the density profile of the core electrons of the individual atoms. The core density would be expected to change much less upon chemical bonding,[33] as can be seen by the fact that the Debye expression with the IAM gives excellent agreement with the essentially exact result at large Q.

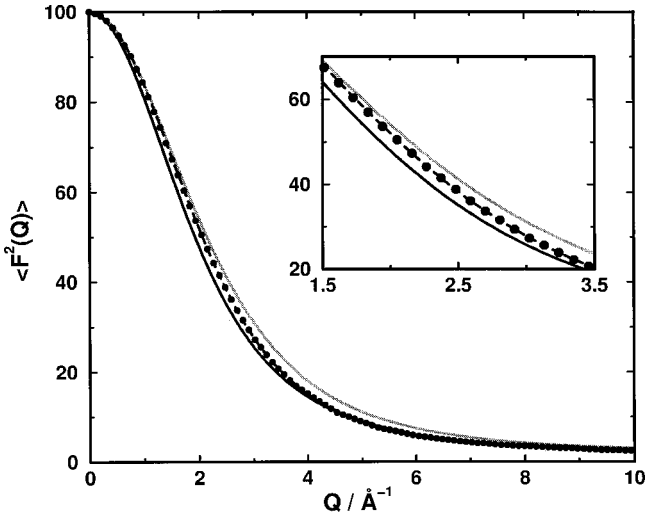

FIG. 2. $\langle F(Q)^2\rangle$ for water in the gas-phase. Legend: Debye with IAM, Eq. (4) (black line); Debye with rescaled atomic scattering factors (gray line); Debye with modified atomic scattering factors, Eq. (5) (dot-dash); four-Gaussian fit to CI calculation (Ref. 26) (circles).

This suggests a modification of the atomic scattering factors which rescales them properly at low Q, but retains their values at large Q. Such a modification is the following:

$$f'(Q)=[1+(\alpha-1)\exp(-Q^2/2\delta^2)]f(Q), \tag{5}$$

where $f'(Q)$ is the modified atomic scattering factor (MASF), $f(Q)$ is the atomic scattering factor for the isolated atom, α is a scaling factor giving the redistribution of charge, and δ is a parameter to be fit, representing the extent of valence-electron delocalization induced by chemical bonding.

For gas phase water, we choose α to correspond to the gas phase dipole moment. The unknown parameter δ can be fit by requiring the Debye expression curve to agree with the *ab initio* configuration interaction (CI) results. In modifying the atomic scattering factors, we could justifiably choose a separate δ for oxygen and hydrogen; however, this introduces a second unknown parameter into our analysis, and we subsequently found this unnecessary. A single parameter choice of $\delta=2.2\,\text{Å}^{-1}$ for both MASFs was found to give excellent agreement when Eq. (5) was inserted into the Debye expression, Eq. (4) (Fig. 2). This single parameter fit to the CI results confirms our chemical intuition and indicates that the zeroth-order change in electron density upon formation of a water molecule from isolated oxygen and hydrogen can be described as a transfer of charge from the hydrogen nuclei to the oxygen nucleus, combined with a greater delocalization of electrons throughout the molecule.

Fitting procedure

The advantage of the MASF formalism lies in the firmer foundation it provides for extraction of the oxygen–oxygen (OO) and possibly oxygen–hydrogen (OH) correlations from the experimental scattering curves. With the proper scaling, they allow the correct weighting of OO and OH correlations, allowing one to extract $g_{\text{OO}}(r)$ and not only a molecular centers radial distribution function. In applying the MASFs to liquid water, we require two parameters for application of Eqs. (2), (4), and (5). α is fixed by the dipole moment of liquid water (that we know lies in the range of 2.6–3.0 D), but δ remains unknown. Lacking knowledge of the true liquid-state $\langle F(Q)^2\rangle$ for water, we will keep δ at its gas-phase value. The sensitivity of our results to both our choice of dipole moment and the valence-electron delocalization parameter will be discussed at the appropriate places below.

Straightforward application of Eqs. (2), (4), and (5) to extract $g_{\text{OO}}(r)$ for water from $h_{\text{OO}}(Q)$ is not possible due to the errors introduced by experimental truncation in Q-space. The $g_{\text{OO}}(r)$ obtained from such a procedure has spurious peaks introduced by the truncation and does not display the proper limiting behavior at small r.[3,31] To reduce these problems, our procedure differed from conventional approaches, such as Q-space continuation[31,36] or minimum noise[5,37] methods, in that we proceeded from combinations of real-space functions to find the optimal $g_{\text{OO}}(r)$ which best fit the experimental data. To find the optimal radial distribution function (RDF) we started with a ''basis set'' of radial distribution functions culled from simulations of various potentials,[6,7,38–41] various experimental curves,[3,5,42] and theoretical predictions.[43] These functions all had the useful properties of displaying the proper small-r behavior, possessing a desirable degree of smoothness, and possessing the proper area under their curves (small-Q behavior).

A linear least-squares program was used to find the optimal coefficients $\{a_i\}$ in the following equation:

$$g_{\text{OO}}(r)=\sum_{i=l}^{N} a_i g^i_{\text{OO}}(r), \tag{6}$$

where the $g_{\text{OO}}(r)$ are the ''basis'' functions taken from the sources described above. Because the $g^i_{\text{OO}}(r)$ described by this equation is a linear superposition of the basis functions, we can perform the optimization entirely in Q-space, using linear superpositions of the corresponding $h^i_{\text{OO}}(Q)$ and resulting in a large savings in computational time.

After an optimal fit has been found from this scheme, various transforms were used to further optimize the curve and overcome any possible basis set limitations. Localized scaling and translation transforms of the form

$$(1+\gamma e^{-\beta(x-x_0)^2})f(x-\epsilon e^{-\lambda(x-x_1)^2})\mapsto f(x) \tag{7}$$

were applied to the optimal $g_{\text{OO}}(r)$ and optimized using Powell's method[44] to further improve the quality of fit. Unlike the above procedure, this process is more computationally intensive because the real-space modifications are no longer linear and a Fourier transform has to be taken at each step to compute the resulting change in scattering. Powell's method was used because derivatives are also no longer available.

The advantages of working in real-space while fitting the experimental curve is that we remain the entire time in a space of functions that satisfy the important small-r and global constraints on possible $g_{\text{OO}}(r)$'s. Truncation errors in Q-space are no longer a problem, and we can address the issue of what range of RDFs will satisfy the data in the given experimental Q-range. However, truncation in r-space is now

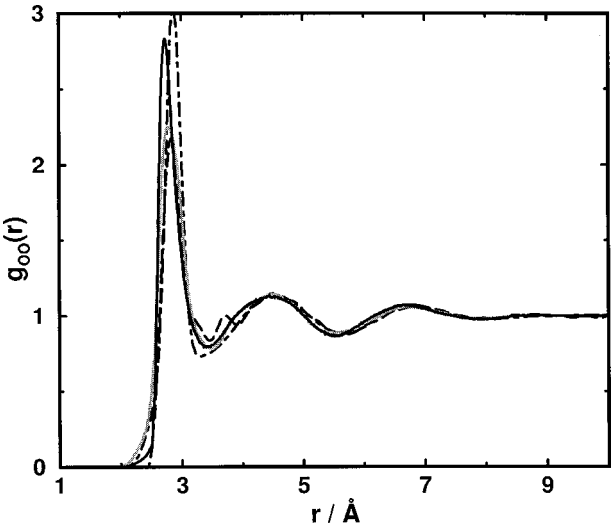

FIG. 3. Comparison of current experimental $g_{OO}(r)$ with previous work. The fit was obtained with $\alpha=1.333$ and $\delta=2.2\,\text{Å}^{-1}$ for the ALS data, x-ray (black line); Narten and Levy, x-ray (Ref. 3) (dashed line); Soper and Philips (Ref. 4) (dot-dashed line). Soper, Bruni, and Ricci, neutron (Ref. 5) (gray line).

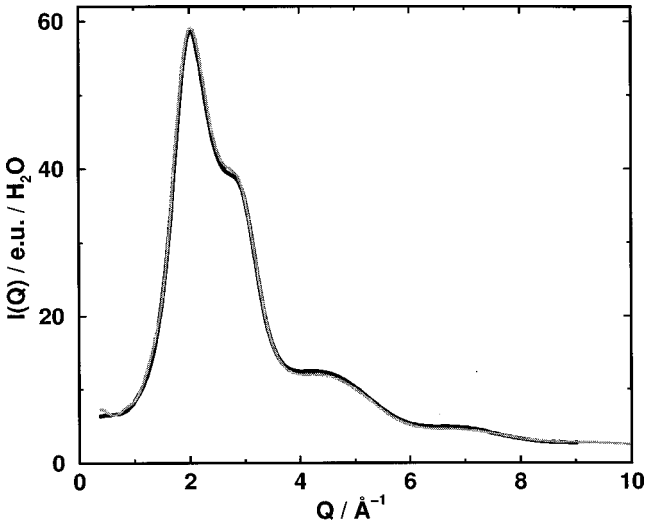

FIG. 4. Predicted x-ray scattering for the $g_{OO}(r)$ in Fig. 3 compared with experiment. Legend: Hura *et al.*, x-ray (black line); best fit (gray line).

a possible problem, depending on how far our basis functions continue in real-space. Fortunately, this only affects the low-Q region of the predicted scattering curve, and this does not affect the quality of our fit since our experimental data is truncated at a relatively high value of $Q=0.4\,\text{Å}^{-1}$. The behavior of the pure water scattering curve at smaller angle is mostly uninteresting and simply describes the lack of significant long-wavelength correlations in liquid water. The $Q=0$ value of the scattering curve gives the isothermal compressibility of water and enforces a constraint on the total integral of the RDFs. However this constraint should be enforced only strictly at the $r=\infty$ limit of the RDFs, and can only be used as an approximate constraint at finite values of r.

The basis set of RDFs used in the present procedure contained values up to $r=12.4\,\text{Å}$, corresponding to simulations of a box of 512 waters; this provides values of predicted scattering up to $Q=0.25\,\text{Å}^{-1}$, sufficient for our present experimental data. With accurate data extending to lower Q, we would need to repeat our procedure with longer RDFs.

RESULTS

To apply the fitting procedure discussed above, we selected various trial functions for $g_{OH}(r)$, and extracted the $g_{OO}(r)$ which best fit the experimental curve with this assumed oxygen–hydrogen correlation. In principle, our fitting procedure could be used to simultaneously fit $g_{OO}(r)$ and $g_{OH}(r)$ to the curve. However, preliminary investigations indicated that the total scattering curve is too insensitive to the form of $g_{OH}(r)$, and a relatively wide range of choices for $g_{OH}(r)$ were consistent within experimental uncertainty and the quality of our fits.

The $g_{OO}(r)$ resulting from this procedure is shown in Fig. 3, assuming the experimental $g_{OH}(r)$ determined by Soper *et al.* from neutron scattering.[5] Also shown in the figure are previous experimental determinations of $g_{OO}(r)$ from Narten's x-ray scattering experiments,[3] the neutron scattering experiment by Soper and Phillips,[4] and recent neutron scattering work by Soper, Bruni, and Ricci.[5] Our fit was obtained using values of $\alpha=1.333$, which corresponds to a liquid-phase dipole moment of 2.8 D, and $\delta=2.2\,\text{Å}^{-1}$, as derived above. The predicted x-ray scattering for this fit is shown in Fig. 4, in comparison to the present experimental scattering curve.

As seen in Fig. 3, the differences between the current fit and earlier determinations of $g_{OO}(r)$ are in the height and sharpness of the first peak, as well as a systematic shift in all peak positions to smaller values of r. What are the meaning of these differences if any? There has been considerable confusion in the simulation community as to what aspects of the radial distribution functions derived from scattering experiments can be quantitatively compared against simulation. The neutron scattering experiments have undergone revision in recent years away from the earlier analysis first presented with the data.[4,5,45] There has also been criticism that the reported experiments by Narten and Levy could not be reproduced based on the information given in Ref. 3.[30] In the development of the TIP4P-FQ model, a comparison of simulated results to those of experiment was considered to the unhelpful,[6] while recent *ab initio* simulations give credence to peak positions of $g_{OO}(r)$ only.[8] In what follows we provide guidance as to what experiment can actually say in regard to this particular measure of ambient water structure.

We first mention that the magnitude of peak and trough heights for the second and third peaks is highly consistent between the very different scattering experiments: recent time of flight neutron scattering data,[5] x-ray diffraction using a reflection geometry setup,[3] and our recently reported ALS data.[18] It is even reasonably consistent between our ALS data and older neutron studies by Soper and Phillips.[4] We would conclude that magnitudes of the second and third peaks are worth reproducing by simulation, and we compare how various water models perform in this regard in the next section.

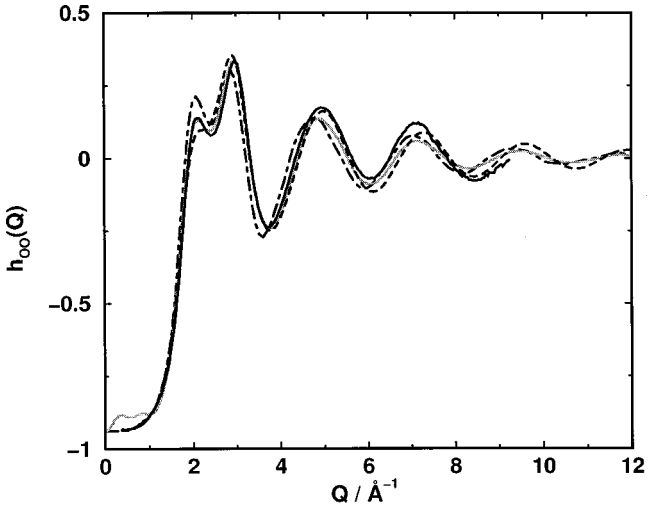

FIG. 5. $h_{OO}(Q)$ from experiment and simulation. Legend: Narten and Levy (Ref. 3) (dot-dash line); Soper, Bruni, and Ricci (Ref. 5) (gray line); current work (black line); SPC/E (Ref. 41) (dashed line). The curve for Narten and Levy is $H_M(Q)$ taken from their article (Ref. 3); the curve for Soper *et al.* is taken from applying a Fourier transform to the $g_{OO}(r)$ given in Ref. 5.

Ironically, given the greater confidence that the scattering experiments should give in regard to defining peak positions accurately, we see that there is still some disagreement between all of the experiments. We know that the primary water peak position of the intensity curve in Q-space for our ALS experiment is shifted to higher angle than older x-ray experiments, a difference that is significant between the two experiments based on error bars estimated from our work.[18] The meaningful shift to higher angle of the intensity is therefore consistent with a determination of a $g_{OO}(r)$ with shifted peak positions to smaller r, and is not an artifact of the fitting procedure. There is a clear trend in the experimental data over the years toward peaks that are shifted to smaller r in $g_{OO}(r)$. We conclude that the quantified errors in our ALS experiment makes a definitive assignment as to the peak position values of $g_{OO}(r)$, and we advocate their reproduction by water simulations to within 1%.

Our analysis shows that a $g_{OO}(r)$ with a higher and sharper first peak is always consistent with the present data, with our best-fit peak height giving a value of 2.8. We find a lower bound to the height of the first peak to be 2.6, which gives an acceptable, but clearly nonoptimal fit to our experimental x-ray intensities. The reported x-ray and neutron studies have reported first peak values of 2.2. Much of the information determining the exact height and shape of this peak is present in the scattering at wave vectors above 7.0 Å^{-1}, and x-ray scattering intensities at much higher angle would be desired to further restrict the peak height.[3,31,36] Nonetheless, we are confident that a higher ($>$2.6) and sharper first peak is correct, and should be a feature captured by empirical water potentials or *ab initio* simulations, but with uncertainty in the peak height value above 2.6.

Proof of this is immediately evident in a comparison of $h_{OO}(Q)$ obtained from the current experiment, Narten and Levy,[3] and Soper, Bruni, and Ricci[5] (Fig. 5). The shorter and broader first peak in $g_{OO}(r)$ reported by the two other experimental groups manifests itself at higher Q as exponentially damped ripples, decaying much faster than the present data. We can extend the ALS data to larger Q in an attempt to extract the $g_{OO}(r)$ straight from this. Figure 5 also shows that simulation models with sharper first peaks such as SPC/E[41] have a much more similar $h_{OO}(Q)$ to our ALS data at large Q, and appear as more logical choices for the extension to even larger values of Q. Doing this, and applying an inverse Fourier transform, retrieves a first peak nearly identical to the first peak present in the SPC/E $g_{OO}(r)$. On the other hand, extending the data by substituting, for example, Narten's $h_{OO}(Q)$ at high-angle produces a $g_{OO}(r)$ with a shortened and broadened first peak. This confirms that this region of $h_{OO}(Q)$ is in part or largely responsible for determining the first peak height and breadth, but also it shows that the data in the range that we have is much more consistent with a taller first peak, rather than showing the heavy exponential damping exhibited in Narten's and Soper's curves.

Interestingly, because of the greater sharpness of the peak, the higher first peak height does not imply a higher coordination number for water. The coordination number is defined as

$$N_c = 4\pi\rho \int_0^{r_{\min}} r^2\, dr\, g_{OO}(r), \tag{8}$$

where ρ is the number density of water and $r_{\min}$ is the location of the first minimum in $g_{OO}(r)$. Application of this formula to the RDFs in Fig. 3 gives values for N_c of 5.1, 5.2, and 4.7, for Narten and Levy,[3,46] Soper, Bruni, and Ricci,[5] and the current work, respectively. A coordination number below five indicates that liquid water preserves much of its ice-like tetrahedral structuring, but with differences in hydrogen-bonding patterns that would also now include deformed or bifurcated hydrogen bonds.[47] An associated liquid such as water should be contrasted to the case of atomic liquids where the coordination number is closer to 12 representing the structure of a fluid dominated by repulsive forces and packing considerations.[48] We see $\sim$10% reduction in N_c from our experiment that, together with greater peak heights and positions shifted to smaller r, suggests liquid water has more ice-like structure than has been the case based on past scattering studies.

Since we propose here a modification of the traditional approach to analyzing x-ray scattering, it was important to investigate the sensitivity of our modeling to the parameters α and δ introduced above. Encouragingly, very similar results for the oxygen–oxygen radial distribution function were obtained for a varying range of α corresponding to dipole moments of 2.6–3.0 D and δ ranging from 1.9 to 2.5 Å^{-1}. This is explainable in view of the fact that the MASF prescription, Eq. (5), changes the atomic scattering factors most in the low-Q region, $\sim$2 Å^{-1} and below, where the resulting impact on the real-space radial distribution function is lower.

Above it was argued that oxygen–hydrogen correlations make only a small contribution to the total x-ray scattering profile, and our fitting procedure has confirmed this to be true. Various $g_{OH}(r)$s drawn from simulation[6,39,40,41] and experiment[5] provide comparable good fits to the observed

curve. The oxygen–oxygen RDF in Fig. 3 was derived by assuming $g_{OH}(r)$ from Ref. 5, but it is interesting that the TIP4P[40] $g_{OH}(r)$ seems to give the closest agreement of any trial oxygen–hydrogen RDF, although the improvement of the fit is not dramatic. The TIP4P RDF mostly differs from Soper's in that the first peak is taller and sharper, further indicating that our analysis finds taller and sharper first peaks consistent with the data. This might be expected since $g(r)$'s extracted with a method which optimizes for smooth functions, such as the minimum noise method used by Soper *et al.*[5] would necessarily favor diminished and broader peaks.

DISCUSSION

The small shift in peak positions and sharper first peak in the oxygen–oxygen radial distribution function, derived by our above analysis on our recent x-ray scattering experiment of liquid water at 27 °C and 1 atm,[18] has several implications for our current understanding of the structure of water. The x-ray scattering profile provides sensitive feedback as to which models are under-structured, possessing too weak of correlations beyond the first solvation shell, and which models are over-structured, possessing too strong of correlations. The predicted scattering profile from simulation of a preferred water model if matched against our experiments can be a good test that the structure of the liquid is being faithfully reproduced. In Figs. 6–9 we compare our $g_{OO}(r)$ derived from our experiment with simulated $g_{OO}(r)$'s using existing nonpolarizable and polarizable empirical force fields for liquid water, as well as recent *ab initio* simulations on liquid water.

Empirical, nonpolarizable water models

Figure 6 shows a comparison of $g_{OO}(r)$ of the ALS experiment with that of SPC,[39] TIP3P,[40] SPC/E,[41] and TIP4P.[40] The first three are three-point charge models while TIP4P is a four-point charge model. All are empirical nonpolarizable water force fields. Based on the comparison of the ALS data in Fig. 6(a), we would argue that SPC and TIP3P are poor descriptions of water structure, especially since tetrahedral structure as measured by the second and third peaks is greatly diminished with respect to experiment.

While the poor structural performance of the SPC and TIP3P water models is appreciated in the simulation community, they are often still used in large size scale or long time scale simulations of pure water or protein–water studies due to the favorable reduced computational cost of simple three-point models. We note that the robust AMBER protein force field was parametrized for use with TIP3P.[49] While a water model should be judged on its global properties, our scattering results discourage the use of models such as SPC and TIP3P. Our recent solution scattering simulations of concentrated N-acetyl-leucine-methylamide in water using AMBER and SPC showed rather poor quantitative agreement with our x-ray scattering experiments.[50,51] A comparison of the simple three-point SPC/E model and the four-point TIP4P nonpolarizable model, show that the overall agreement of these models with experiment is quite good.

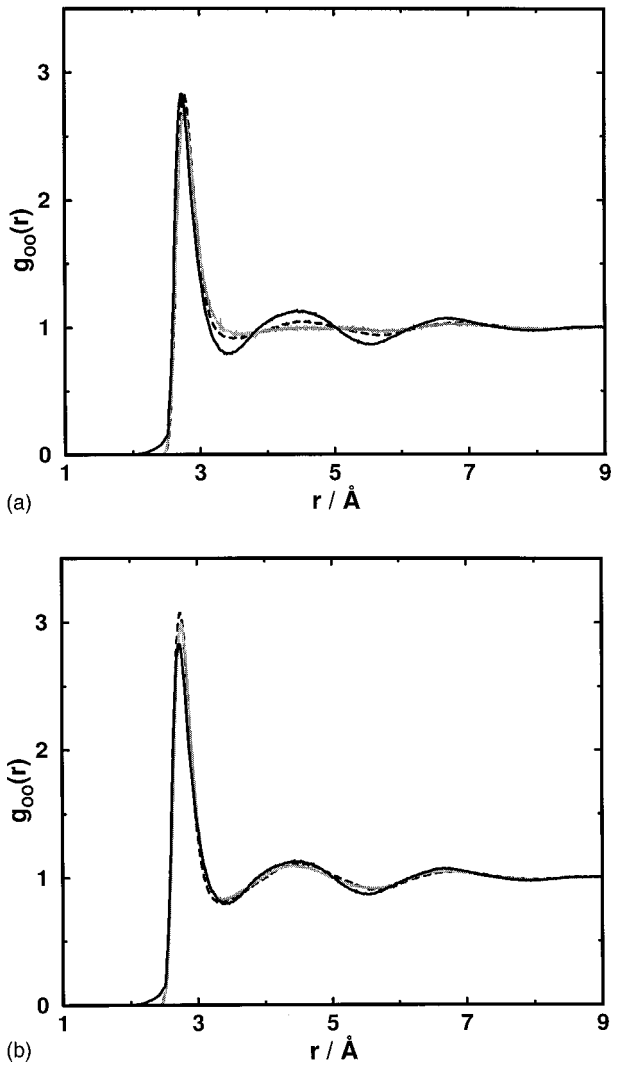

FIG. 6. Comparison of current experimental $g_{OO}(r)$ (black line) with simulations using nonpolarizable water models. (a) TIP3P (gray line) and SPC (dashed line) three-point models do a poor job of reproducing the ALS data. (b) TIP4P (gray line) and SPC/E (dashed line), four- and three-point models, show better agreement with the ALS data.

Figure 7 displays a comparison of the ALS $g_{OO}(r)$ with that of the five-point ST2,[52] ST4,[7] and and recently proposed TIP5P[53] water models. ST2 was the first molecular water model developed to provide a simulation-based description of water over a wide range of conditions, and is historically significant because qualitative insight into the diverse molecular properties of water were successfully obtained.[52] ST2 has been criticized as being ''too tetrahedral,'' [9,40] and it remains the case that, when compared to our ALS-derived $g_{OO}(r)$, ST2 is over-structured. ST4 is a modification of ST2 that was pursued to study the orientational effects in the liquid,[7] and was designed to reproduce the neutron scattering $g_{OO}(r)$ reported by Soper and Phillips (SP).[4] When comparing the ALS data against ST4 simulation, we see some noticeable structural improvement for ST4 relative to ST2, but it is still inadequate, and by association so is the older neutron study of water.[4] TIP5P is a new water model parametrized independent of our ALS data that reproduces the density of liquid water accurately over a temperature range from

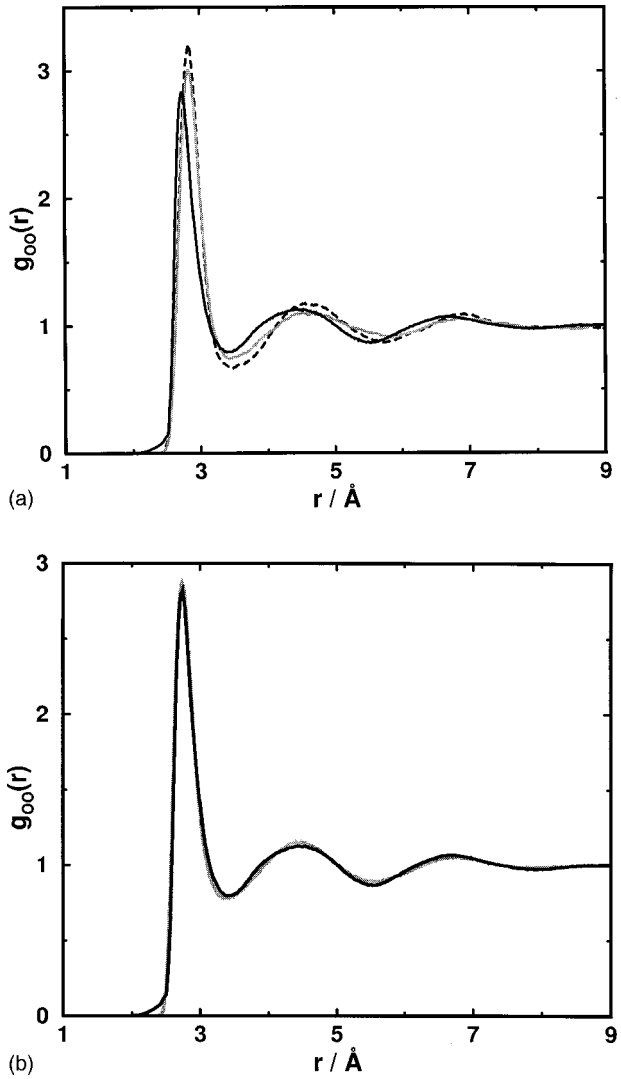

FIG. 7. Comparison of current experimental $g_{OO}(r)$ (black line) with simulations of five-point nonpolarizable water models. (a) ST2 (dashed line) and ST4 (gray line); ST4 is a variant of ST2 that was designed to reproduce accurately the original neutron experimental data from Soper and Phillips (Ref. 4). (b) Comparison to the newest water model TIP5P (gray line) that shows the best agreement compared to all simulations of liquid water examined in this work.

$-37.5\,^{\circ}\mathrm{C}$ to $62.5\,^{\circ}\mathrm{C}$ at 1 atm.[53] The agreement of the TIP5P simulation with the ALS data is remarkable, and as far as $g_{OO}(r)$ is concerned, is a noteworthy improvement over TIP4P.

Polarizable water models

While the nonpolarizable models SPC/E and TIP4P show good agreement with our ALS data, and the new TIP5P model gives excellent agreement, applications of nonpolarizable models to water phenomena away from ambient conditions[5,45] or aqueous solution structural studies[50,51] have proven them to be deficient when taken outside their optimal parametrized state. We expect that the next generation of empirical force fields will include polarizability, often accepted as a necessary means of improving quantitative agreement between simulations and experiments away from ambient conditions or for heterogeneous chemical systems.[45,54,55]

Figure 8 shows a comparison of our ALS $g_{OO}(r)$ with several polarizable water models including a fluctuating charge version of TIP4P that we call TIP4P-FQ,[6] an extension of TIP4P-FQ that introduces an additional coupling between the Lennard-Jones interaction parameters for a pair of oxygen sites and their partial charges,[56] TIP4P-Pol-1, extension of the MCY water model to include flexible bonds and angles, as well as many-body effects, NCC-vib,[57] the polarizable point charge (PPC) model,[58] and a simple polarizable model developed to reproduce water properties over a wide range of conditions by Chialvo and Cummings (CC).[45]

The CC model shifts all peaks to larger *r*, and has a very large peak as well as overemphasis on the loss of density at the first minimum.[59] While overall its structure does least well among the polarizable water models, its parametrization may well reproduce nonambient states better.[45] The NCC-vib model[57] also overemphasizes the loss of density at the first minimum, but is otherwise faithful to the experimentally determined $g_{OO}(r)$. For the TIP4P-FQ model[6] and the PPC model,[58] it is evident that the overall agreement is excellent, although the position of the experimental first peak is not as well-reproduced as the nonpolarizable TIP5P model. The TIP4P-pol-1 water model shows improvements in first peak positions relative to TIP4P-FQ, but overemphasizes loss and gain of density at the first minimum and second maximum, respectively. While overall the polarizable models perform well, these new generation of force fields are not optimal performers at ambient temperature as are some of their nonpolarizable partners.

Ab initio molecular dynamics simulations

The first fully quantum treatments of electronic structure in aqueous simulations are beginning to emerge,[22,60–63] and provide another interesting point of comparison to our experiments and theoretical analysis. All rely on a local density approximation to density functional theory. Differences among functionals in the condensed phase have been investigated for liquid water in Ref. 60, and showed that BLYP[64,65] performed best compared to other density functionals considered at that time. The more recently introduced PBE[66] functional used in the *ab initio* simulation reported in Ref. 62 gives binding energies and oxygen–oxygen distances that are closer to a large basis set MP2 calculation[67] on the water dimer relative to the BLYP functional.

In Figure 9 we show a comparison of our ALS-derived data with several recently reported *ab initio* simulated $g_{OO}(r)$'s. These include a 10 ps, 64 water molecule molecular dynamics (MD) run with a gradient-corrected BLYP functional with an average ionic temperature of 318 K,[8] a 5 ps, 32 water molecule MD run with the gradient-corrected BLYP functional and average ionic temperature of 303 K,[61] a 2 ps, 54 water molecule MD run with the PBE functional, and at a temperature of ~300 K,[62] and a new 6.4 ps *ab initio* simulation with the PBE functional at ~294 K by Schwegler and co-workers,[63] to be reported in a future publication.

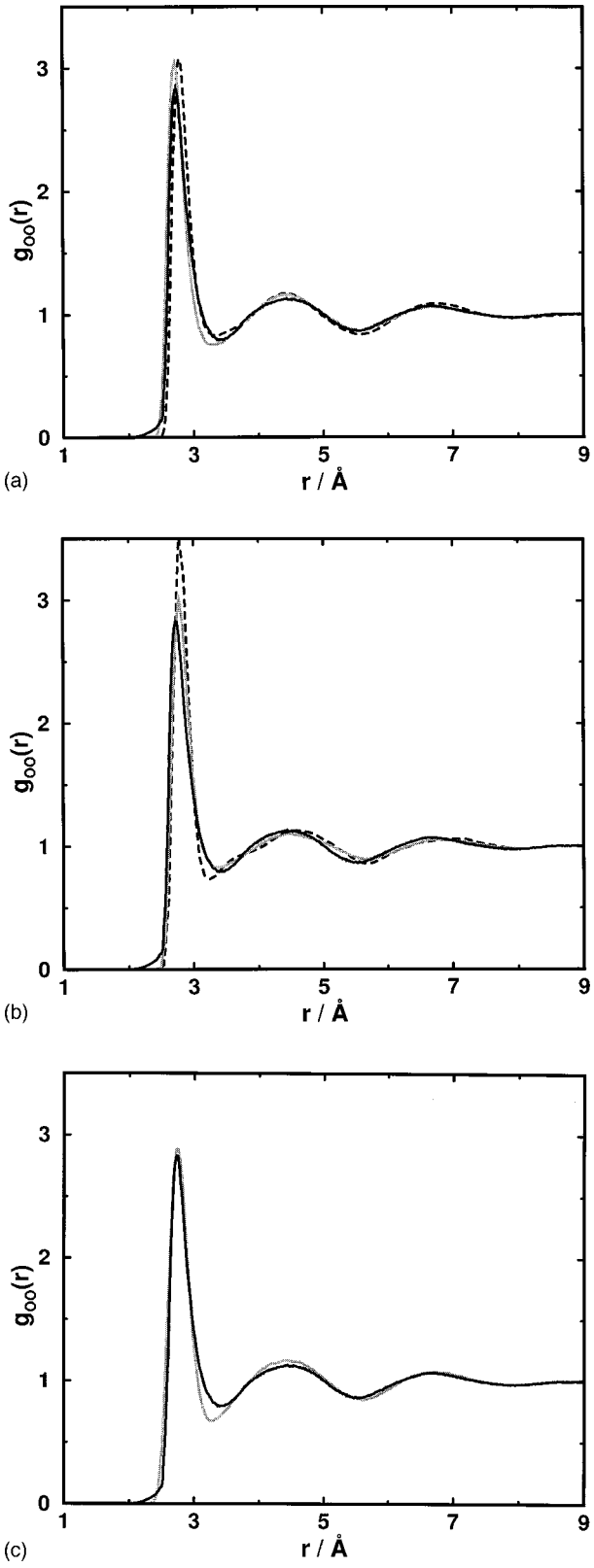

FIG. 8. Comparison of current experimental $g_{OO}(r)$ (black line) with simulations using polarizable water models. (a) The TIP4P-Pol-1 model (Ref. 56) (gray line) and TIP4P-FQ model (Ref. 6) (dashed line). (b) The CC model (Ref. 45) (dashed line) and the PPC model (Ref. 58) (gray line). (c) The NCC-Vib model (gray line) (Ref. 57).

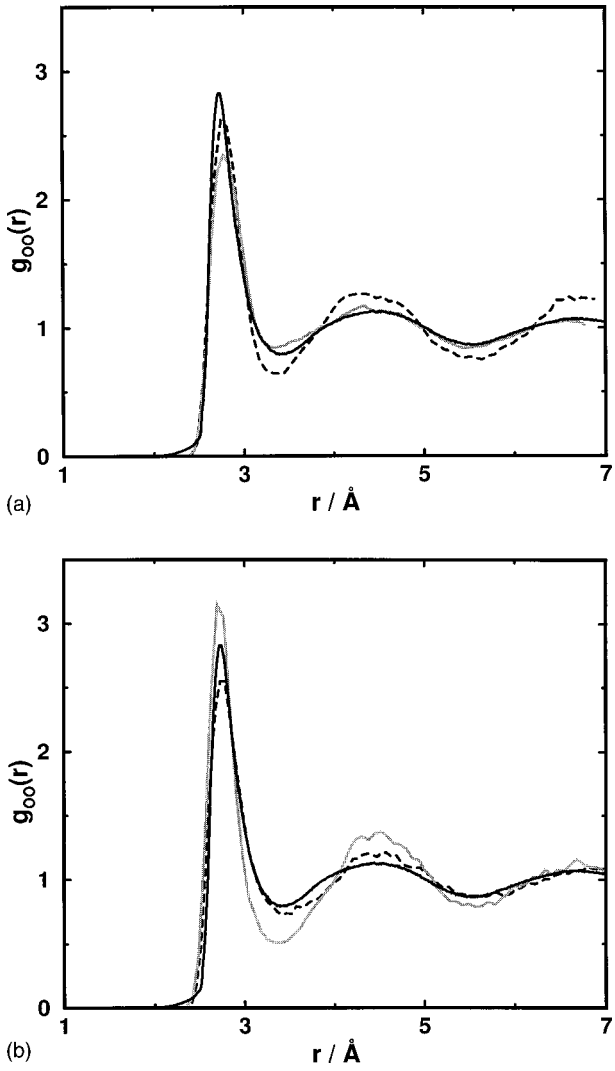

FIG. 9. Comparison of ALS experimental $g_{OO}(r)$ (black line) with *ab initio* molecular dynamics simulations. (a) Silvestrelli and Parrinello (Ref. 8) *ab initio* simulation of 10 ps for 64 water molecules, average ionic temperature of 318 K (gray line), Sprik *et al.* (Ref. 61) *ab initio* simulation of 5 ps for 32 water molecules, average ionic temperature of 303 K (dashed line), (b) Schwegler *et al.* (Ref. 62) *ab initio* simulation of 2 ps for 54 water molecule average ionic temperature of ~300 K (gray line); more recent *ab initio* simulation by Schwegler *et al.* (Ref. 63) 3 ps for 54 water molecules, average ionic temperature of ~294 K (dashed line).

While the quantitative agreement is not as adequate as one would hope, we will see that some aspects of the comparison are favorable.

In the case of the published *ab initio* simulation using the PBE functional,[62] the reported $g_{OO}(r)$ was based on an initial water configuration taken from a TIP4P simulation, first simulated at $T=600\,\mathrm{K}$, $\rho=0.9\,\mathrm{g/cc}$, then cooled to $T=300\,\mathrm{K}$ at the same density, and finally the density was increased to $\rho=1.0\,\mathrm{g/cc}$ at room temperature. In each case of a given density and temperature point, the system was equilibrated with a thermostat for approximately 0.5 ps (this does not include the time it took to cool or heat the simulation), and then data was collected with the thermostat off for 2.0 ps. The *ab initio* simulations using the PBE functional were much shorter than other BLYP *ab initio* simulations for several reasons. First H_2O and not D_2O was simulated, so that

the time step used was three times smaller than other studies (~0.05 fs). Second this calculation used a plane wave cutoff of the pseudopotential for oxygen of 80 Ry, which is necessary to reliably calculate pressure in water, which was the purpose of the PBE study.[62]

The positive aspects of the PBE simulation reported in Ref. 62 is that it gets peak positions spot on, and the distorted magnitudes and shape of $g_{OO}(r)$ might improve with longer simulation runs. Currently, however, we suspect that the system had not reached an equilibrated state in this short amount of simulation time, and therefore that statistics collected over 2 ps have not given converged results. We think the reported ambient $g_{OO}(r)$ from this simulation shows residual structure from its ''history,'' i.e., the 0.9 g/cc low density state that is also known to give greater magnitudes in peak positions that arise from local regions of relatively high density.[68] This was partly confirmed by a more recent *ab initio* simulation that was started from the simulation reported in Ref. 62 (final temperature and density of 300 K, 1.0 g/cc), heated to 600 K then cooled to $T=300\,\mathrm{K}$, and at all stages maintaining a density of 1.0 g/cc. The average ionic temperature is 294 K. The 6.4 ps simulation results shown in Fig. 9(b) show much improvement relative to the original data reported in Ref. 62, and will be analyzed further in Ref. 63.

In the case of the *ab initio* simulations using the gradient-corrected BLYP functional,[61] the originally reported $g_{OO}(r)$ was derived from a simulation that was based on a 32 water molecule configuration taken from a TIP4P classical simulation. The system was equilibrated with a thermostat for approximately 0.5 ps, and then data was collected with the thermostat off for 5 ps with an average ionic temperature of 303 K. More recently, an *ab initio* simulation[8] with 64 water molecules reports a new $g_{OO}(r)$ derived from a configuration taken from a body-centered cubic lattice with random orientations, equilibrated with a thermostat for an unreported length of time, and then data was collected with the thermostat off for 10 ps, giving an average ionic temperature of 318 K.

The comparison of the 32 water and 64 water molecule BLYP functional simulations to our ALS data as well as past experiments,[3–5] is complicated by both finite size effects and higher temperatures, and it is worthwhile discussing these well-known effects on structure that can be found in the classical simulation literature.[69–72] We show in Fig. 10 a simulation with the TIP4P-FQ model for 512 waters at two different temperatures: 298 K and at 318 K [Fig. 10(a)] as well as differences in $g_{OO}(r)$ due to finite size effects by considering 32, 64, and 512 waters at 298 K [Fig. 10(b)]. We chose the polarizable water model since the interaction potential complexity mimics the electronic distortion seen in an *ab initio* simulation, unlike the nonpolarizable water models, giving us a better idea how temperature and finite size effects may influence the *ab initio* results. As we see from Fig. 10(a), higher temperatures result in an overall softening of the structure, with diminishment of the first and second peak, and a rise in the first minimum. As seen in Fig. 10(b) the structural distortions due to finite size effects are significant for 32 water molecules, but largely disappear for system

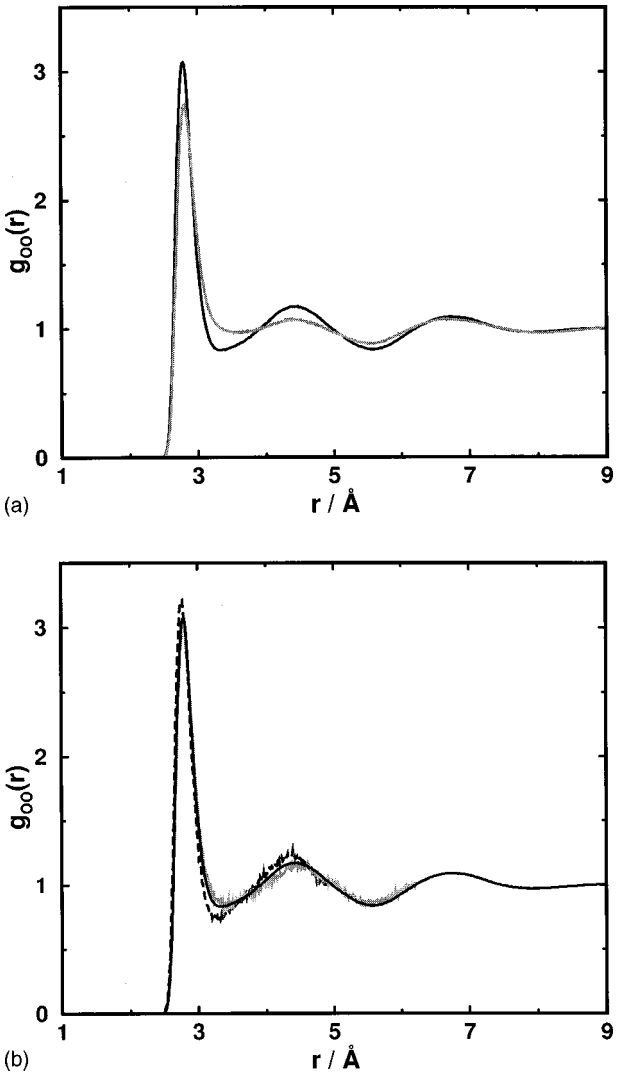

FIG. 10. The changes in $g_{OO}(r)$ expected as a function of (a) temperature, 298 K (black line) and 318 K (gray line) and (b) finite size effects for 32 waters (dashed line), 64 waters (gray line), and 512 waters (black line), using the polarizable TIP-FQ model (Ref. 6).

sizes of ~64 or more water molecules. This suggests that the smaller system *ab initio* simulations would be somewhat over-structured, with peak positions shifted to smaller *r*. Finite size effects are clearly apparent for 32 water molecule simulations with BLYP,[61] and therefore some of the structural differences seen in those simulations are clearly due to the small box size used. In the case of the 64 water molecule simulation with BLYP[8] the finite size effects are greatly diminished (as is probably the case for 54 water molecules for the BP simulation).[62] However, the higher temperature classical simulation shows structural differences at room temperature that are partly consistent with changes observed between the *ab initio* simulations at ~300 K and 318 K, namely in the first peak, and some population in the first minimum. It is not immediately clear that differences cited between the 32 and 64 water molecule simulations are largely due to finite size effects,[8] but instead we would also attribute the observed structural changes to be due to the higher temperature.

CONCLUSION

We have presented an analysis of our recent ALS x-ray experiment on pure liquid water at ambient temperature and pressure. The traditional application of Eq. (2) to analysis of x-ray scattering employs the atomic scattering factors derived from isolated atom calculations.[20] With the current experimental results, we found that using these atomic scattering factors gave a much less satisfactory fit to the data. An important part of this present work is to advocate the development of better methods for quantitatively analyzing liquid-state x-ray scattering. Both the condensed phase form factor and the atomic scattering factors will always have to come from models, preferably derived from condensed-phase calculations. It seems that the ideal approach would be to do away with this level of modeling and the allocation of electrons into convenient parts and to instead apply Eq. (1) without further assumptions. With the advent of completely *ab initio* simulations of liquid water,[8,60–63] this would now seem possible. As well as publishing radial distribution functions probing nuclear–nuclear correlations, *ab initio* simulations could predict the x-ray scattering straight from the distribution of electron density in their simulation box. This would ultimately provide a much more stringent test of the current *ab initio* simulations and give more confidence in *ab initio* predictions of the water radial distribution functions.

We have made satisfactory progress using Eq. (2), however, where we have argued that considerations of the charge density of the water molecule in the condensed phase requires a modification of the isolated atom scattering factors; this difference proved sufficient to give a much stronger fit for the analysis using the MASF model. Furthermore, it is clear based on this analysis that an x-ray scattering experiment largely contains information about $g_{OO}(r)$, with only a small component due to the O–H correlations. This means that our reported measurement will be more reliable for OO correlations than that determined by neutron scattering, and the quality of our experiment make it preferable to that of the original Narten and Levy data.

We presented a $g_{OO}(r)$ for water consistent with our recent experimental data gathered at the ALS, which is different in some aspects than the $g_{OO}(r)$ reported by other x-ray and neutron scattering experiments. We find a taller and sharper first peak, and systematic shifts in all peak positions to smaller r, relative to past reported experimental $g_{OO}(r)$ data. Because much of the experimental information on the first peak in $g_{OO}(r)$ resides in the tail of the data at larger values of Q than probed by the current experiment, we cannot precisely assert the exact height of the $g_{OO}(r)$ peak other than to say that it should be greater than 2.6. We do find, however, consistent agreement on second and third peak magnitudes between all experiments which is worthwhile reproducing by the simulation community. Based on our experiments discussed in the companion article, and theoretical analysis in this work, we would argue that liquid water is more structured than that determined from past scattering experiments.

Given the current experimental picture concerning the structure of liquid water as probed by x-ray and neutron scattering experiments, we have outlined what features of $g_{OO}(r)$ should be reproduced by a water model, and which water models favorably reproduce those features. Table I provides a structural summary of peak and trough positions, and their magnitudes for various nonpolarizable, polarizable, *ab initio* simulations, and experiment. In the nonpolarizable water model category we find that TIP3P, SPC, ST2, ST4, as well as MCY (data not shown) are inadequate structural descriptions of ambient water, while SPC/E and TIP4P give good agreement with our ALS experiment. The recently introduced TIP5P five-site model[53] gives excellent agreement with our ALS data, and in addition to its robust performance in accurate densities over a large temperature range, makes it possibly the current nonpolarizable water model of choice in classical simulation. We find that the NCC-vib, PPC, and TIP4P-FQ, and TIP4P-Pol-1 polarizable models perform quite well, but with some problems in the vicinity of the first peak. The CC model has similar problems to the former polarizable models, but also shows shifts in all peak positions to larger r than what we determine from experiment.

The problems with the reported *ab initio* results arise from several sources that typically have been investigated and overcome in the classical simulation literature.[68–72] These include dependence on initial conditions, length of the simulations, variation in system properties that arise with temperature[70] or density,[68] and finite size effects.[71,72] First

TABLE I. $g_{OO}(r)$ of various nonpolarizable and polarizable effective potential models of water, *ab initio* simulations, and experiments. Data is listed as r, $g_{OO}(r)$ pairs.

$g_{OO}(r)$	1st peak	1st minimum	2nd peak	2nd minimum	3rd peak
SPC[a]	2.77, 2.82	3.49, 0.91	4.53, 1.04	5.70, 0.93	6.85, 1.04
SPC[b]	2.75, 3.07	3.31, 0.80	4.49, 1.11	5.65, 0.90	6.83, 1.05
TIP3P[c]	2.77, 2.69	3.53, 0.94	4.63, 1.00	5.77, 0.96	6.81, 1.03
TIP4P[c]	2.76, 2.98	3.36, 0.82	4.40, 1.09	5.60, 0.91	6.71, 1.05
TIP5P[d]	2.73, 2.87	3.37, 0.78	4.47, 1.15	5.60, 0.89	6.72, 1.05
TIP4P-FQ[e]	2.79, 3.07	3.32, 0.83	4.43, 1.17	5.58, 0.84	6.73, 1.09
CC[f]	2.77, 3.52	3.25, 0.73	4.60, 1.13	5.66, 0.86	6.99, 1.06
NCC-vib[g]	2.74, 2.89	3.27, 0.67	4.41, 1.17	5.61, 0.84	6.71, 1.08
PPC[h]	2.77, 3.04	3.38, 0.82	4.48, 1.10	5.65, 0.89	6.81, 1.05
TIP4P-Pol-1[i]	2.71, 3.04	3.31, 0.75	4.37, 1.16	5.56, 0.87	6.67, 1.06
ST2[j]	2.83, 3.20	3.48, 0.67	4.61, 1.18	5.71, 0.86	6.91, 1.09
ST4[k]	2.83, 3.00	3.43, 0.74	4.59, 1.10	5.82, 0.90	6.89, 1.05
CPMD (32)[l]	2.76, 2.62	3.37, 0.64	4.31, 1.27	5.55, 0.75	6.68, 1.24
CPMD (64)[m]	2.78, 2.35	3.32, 0.83	4.39, 1.17	5.47, 0.84	6.56, 1.06
LLNL (54)[n]	2.71, 3.17	3.36, 0.50	4.44, 1.37	5.56, 0.78	6.69, 1.16
New LLNL[o]	2.78, 2.60	3.47, 0.69	4.52, 1.20	5.58, 0.88	6.80, 1.12
Narten and Levy[p]	2.84, 2.18	3.45, 0.83	4.50, 1.15	5.55, 0.87	6.86, 1.07
Soper and Phillips[q]	2.88, 3.09	3.33, 0.73	4.50, 1.14	5.68, 0.88	6.75, 1.07
Soper *et al.*[r]	2.79, 2.25	3.49, 0.78	4.48, 1.15	5.51, 0.89	6.72, 1.06
ALS[s]	2.73, 2.83	3.41, 0.79	4.44, 1.13	5.51, 0.86	6.66, 1.07

[a]Reference 39.
[b]Reference 41.
[c]Reference 40.
[d]Reference 53.
[e]Reference 6.
[f]Reference 45.
[g]Reference 57.
[h]Reference 58.
[i]Reference 56.
[j]Reference 52.
[k]Reference 7.
[l]Reference 61.
[m]Reference 8.
[n]Reference 62.
[o]Reference 63.
[p]Reference 3.
[q]Reference 4.
[r]Reference 5.
[s]Reference 18.

principles approaches to condensed phase simulations can be improved in a number of ways. A minimum of 10 ps is likely required to report meaningful averages, since the time required for water to diffuse one molecular diameter is 6 ps based on the experimental diffusion constant of 2.4×10^{-5} cm^2/s. Because temperature, pressure, and density effects can have significant effects on structure, it is important to compare the same thermodynamic state between experiment and simulation. It is also interesting to note that the transferability of a functional designed with gas-phase properties in mind will need proper calibration of its performance in the condensed phase (as was done in Ref. 60), since BLYP and PBE do not include long-ranged dispersion. Given the current computational expense that prohibit box sizes typically used in empirical force field simulations at present, these are largely technical limitations that will clearly diminish over time, and we would expect quantitative agreement to improve in the future.

It is interesting to consider what are the underlying goals of the classical and *ab initio* approaches to condensed phase simulation of systems such as liquid water. In the classical simulation arena, fixed charge models can be viewed as ''effective'' potentials that inherently include quantum effects and electronic polarization through parametrized condensed phase values of the water dipole moment. Polarizable empirical force fields fall into the same category of effective potentials, although they treat many-body effects in a more explicit way. The next generation of empirical force fields that will include polarizability is often motivated as a necessary means of improving quantitative agreement between simulations and experiments away from ambient conditions or for heterogeneous chemical systems.[45,54,55] The striking agreement of the five-site, fixed charge TIP5P water model with our ALS-derived $g_{OO}(r)$, and its ability to reproduce water properties over a good range of temperatures and pressures, may challenge this assumption in the future for pure water. However, both nonpolarizable and polarizable force fields should be tested for their performance on solution studies where mixing independently derived water, and protein force fields can give rise to inadequate structure and energetics.

Overlaying quantum corrections on these effective classical potential models of water, known to ''soften'' the structure of a simulated water model,[73–75] would improve some classical models and diminish agreement for others with respect to our experiment and analysis. It would seem that the addition of quantum corrections is incompatible with the philosophy of classical force fields, unlike the philosophy underpinning first principles approaches to molecular dynamics simulations in the condensed phase. In addition to adequate simulation cell sizes, and sufficient time to equilibrate and collect statistics, the *ab initio* simulation may also require the addition of quantum corrections using *ab initio* path integral simulations to make a comparison at a given thermodynamic state complete.[76] Whether *ab initio* simulations using the PBE and BLYP functionals would improve with complete quantum modeling remains to be investigated, although the higher temperature BLYP simulation of water by Silvestrelli and Parrinello[8] provide preliminary evidence for quantitative improvement in this regard.

It has been noted that the parametrization of the global properties of water, i.e., developing a water model that exhibits reasonable dielectric and diffusion constants, pressure, density maximum, in addition to sensible structure, is inconsistent with a model that gives a shorter and more broad first peak of the oxygen–oxygen radial distribution function.[77] It is interesting to note that almost all empirical water models, regardless of the number of charge centers and whether polarizable or nonpolarizable, exhibit a relatively high and sharp first peak, contrary to past experiments. In fact, with some few exceptions such as ST4[7] and a water model developed by Watanabe and Klein,[9] most water models have not been parametrized to reproduce the high first peak of the original SP data as was stated in Ref. 8.

In conclusion, extension of our experiments, simulations, and theoretical analysis will include investigations at higher and lower temperatures and pressures to aid in the development of water models and simulations that are successful over a wide range of thermodynamic states. We have just completed a series of runs at the ALS on liquid water over a range of temperatures of 10 °C to 70 °C at 1 atm, that we will report on in the near future.[78] In our own group, we require quantitative improvements of aqueous solution simulations to use with solution scattering experiments[50,51] to investigate hydration forces in protein folding and stability, and to probe how protein folding, stability, and kinetics change with temperature and pressure.

ACKNOWLEDGMENTS

J.M.S. is supported by a National Science Foundation Graduate Research fellowship. G.H. and T.H.G. thank the Office of Biological and Environmental Research (OBER), U.S. Department of Energy Contract No. DE-AC-03-76SF0098 for support in FY99. T.H.G. gratefully acknowledges support from the Air Force Office of Sponsored Research, Grant No. FQ8671-9601129, and the National Energy Research Supercomputer Center for computer time. We would like to thank the following for contributing to us their tabulated oxygen–oxygen radial distribution function and comments on the manuscript: E. Schwegler, G. Galli, P. L. Silvestrelli, A. Chialvo, G. Corongiu, I. M. Svishchev, P. G. Kusalik, R. J. Boyd, M. Mahoney, W. Jorgensen, and J. I. Siepmann. We also thank F. Stillinger for discussions involving water model parametrization.

[1] W. C. Roentgen, Annu. Rev. Phys. Chem. **45**, 91 (1892).
[2] J. D. Bernal and R. H. Fowler, J. Chem. Phys. **1**, 515 (1933).
[3] A. H. Narten and H. A. Levy, J. Chem. Phys. **55**, 2263 (1971).
[4] A. K. Soper and M. G. Phillips, Chem. Phys. **107**, 47 (1986).
[5] A. K. Soper, F. Bruni, and M. A. Ricci, J. Chem. Phys. **106**, 247 (1997).
[6] Y.-P. Liu, K. Kim, B. J. Berne, R. A. Friesner, and S. W. Rick, J. Chem. Phys. **108**, 4739 (1998).
[7] T. Head-Gordon and F. H. Stillinger, J. Chem. Phys. **98**, 3313 (1993).
[8] P. L. Silvestrelli and M. Parrinello, J. Chem. Phys. **111**, 3572 (1999).
[9] K. Watanabe and M. L. Klein, Chem. Phys. **131**, 157 (1989).
[10] T. Lazaridis and M. E. Paulaitis, J. Phys. Chem. **96**, 3847 (1992).
[11] K. Ding, D. Chandler, S. J. Smithline, and A. D. J. Haymet, Phys. Rev. Lett. **59**, 1698 (1987).

[12] G. Hummer, S. Garde, A. E. Garcia *et al.*, J. Phys. Chem. B **102**, 10469 (1998).
[13] L. R. Pratt and D. Chandler, J. Chem. Phys. **67**, 3683 (1977).
[14] T. Head-Gordon, J. Am. Chem. Soc. **117**, 501 (1995).
[15] F. H. Stillinger, J. Solution Chem. **2**, 141 (1973).
[16] T. Lazaridis and M. Karplus, J. Chem. Phys. **105**, 4294 (1996).
[17] N. Matubayasi, E. Gallicchio, and R. M. Levy, J. Chem. Phys. **109**, 4864 (1998).
[18] G. Hura, J. Sorenson, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9140 (2000), preceding paper.
[19] P. A. Egelstaff, *An Introduction to the Liquid State*, 2nd ed. (Clarendon, Oxford, 1992).
[20] *International Table for X-ray Crystallography*, 3rd ed., edited by J. A. Ibers and W. C. Hamilton (Kluwer Academic, Dordrecht, 1989), Vol. 3.
[21] A. Guinier, *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies* (Dover, New York, 1994).
[22] L. Blum and A. H. Narten, Adv. Chem. Phys. **34**, 203 (1976).
[23] K. Hermansson, Chem. Phys. Lett. **260**, 229 (1996).
[24] L. Blum, J. Comput. Phys. **7**, 592 (1971).
[25] R. Moccia, J. Chem. Phys. **40**, 2186 (1964).
[26] J. Wang, A. N. Tripathi, and V. H. Smith, Jr., J. Chem. Phys. **101**, 4842 (1994).
[27] H. Takeuchi, M. Nakagawa, T. Saito, T. Egawa, K. Tanaka, and S. Konaka, Int. J. Quantum Chem. **52**, 1339 (1994).
[28] S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, J. Chem. Phys. **59**, 2254 (1973).
[29] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).
[30] P. A. Egelstaff and J. H. Root, Chem. Phys. **76**, 405 (1983).
[31] A. V. Okhulkov, Y. N. Demianets, and Y. E. Gorbaty, J. Chem. Phys. **100**, 1578 (1994).
[32] E. R. Batista, S. S. Xantheas, and H. J'onsson, J. Chem. Phys. **109**, 4546 (1998).
[33] P. Coppens, *X-ray Charge Densities and Chemical Bonding* (Oxford University Press, Oxford, 1997).
[34] P. Coppens, T. N. Guru-Row, P. Leung, E. D. Stevens, P. J. Becker, and Y. W. Yang, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **35**, 63 (1979).
[35] N. K. Hansen and P. Coppens, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **34**, 909 (1978).
[36] A. H. Narten, C. G. Venkatesh, and S. A. Rice, J. Chem. Phys. **64**, 1106 (1976).
[37] A. K. Soper, C. Andreani, and M. Nardone, Phys. Rev. E **47**, 2598 (1993).
[38] G. C. Lie, E. Clementi, and M. Yoshimine, J. Chem. Phys. **64**, 2314 (1976).
[39] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), p. 331.
[40] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[41] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[42] Y. E. Gorbaty and Y. N. Demianets, Mol. Phys. **55**, 571 (1985).
[43] B. M. Pettitt and P. J. Rossky, J. Chem. Phys. **77**, 1451 (1982).
[44] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
[45] A. A. Chialvo and P. T. Cummings, J. Chem. Phys. **105**, 8274 (1996).
[46] Application of Eq. (7) to Narten's data depends sensitively on the location of $r_{\min}$. The value of the running coordination number we give in the text is for a choice of $r_{\min}=3.6$.
[47] A. Rahman and F. H. Stillinger, J. Am. Chem. Soc. **95**, 7943 (1973).
[48] D. Chandler, J. D. Weeks, and H. C. Anderson, Science **220**, 787 (1983).
[49] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[50] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, Perspect. Drug Discovery Des. **17**, 97 (1999).
[51] J. M. Sorenson, G. Hura, A. K. Soper, A. Pertsemlidis, and T. Head-Gordon, J. Phys. Chem. B **103**, 5413 (1999).
[52] A. Rahman and F. H. Stillinger, J. Chem. Phys. **55**, 3336 (1971).
[53] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[54] T.-M. Chang and L. X. Dang, J. Chem. Phys. **104**, 6772 (1996).
[55] S. W. Rick, S. J. Stuart, and B. J. Berne, J. Chem. Phys. **101**, 6141 (1994).
[56] B. Chen, J. Xing and J. I. Siepmann, J. Phys. Chem. B **104**, 2391 (2000).
[57] G. Corongiu and E. Clementi, J. Phys. Chem. **97**, 2030 (1992).
[58] I. M. Svishchev, P. G. Kusalik, J. Wang, and R. J. Boyd, J. Chem. Phys. **105**, 4742 (1996).
[59] We are using the $R_{\mathrm{OM}}=0.2$ Å version of the CC model described in Ref. 45.
[60] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).
[61] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. **105**, 1142 (1996).
[62] E. Schwegler, G. Galli, and F. Gygi, Phys. Rev. Lett. **84**, 2429 (2000).
[63] E. Schwegler and G. Galli (unpublished).
[64] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[65] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[66] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[67] I. M. B. Nielsen, E. T. Seidl, and C. L. Janssen, J. Chem. Phys. **110**, 9435 (1999).
[68] M. E. Parker and D. M. Heyes, J. Chem. Phys. **108**, 9039 (1998).
[69] M. P. Allen and D. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[70] L. A. Báez and P. Clancy, J. Chem. Phys. **101**, 9837 (1994).
[71] A. R. Denton and P. A. Egelstaff, Z. Phys. B: Condens. Matter **103**, 343 (1997).
[72] L. R. Pratt and S. W. Haan, J. Chem. Phys. **74**, 1864 (1981).
[73] G. S. Del Buono, P. J. Rossky, and J. Schnitker, J. Chem. Phys. **95**, 3728 (1991).
[74] J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).
[75] B. Guillot and Y. Guissani, J. Chem. Phys. **108**, 10162 (1998).
[76] D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, Nature (London) **397**, 601 (1999).
[77] F. H. Stillinger (private communication).
[78] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon (unpublished).